超越复活节岛陷阱的太空时代，生命和智力演化会大大加快。未来大学通识教育应基于科学逻辑并解释智力科学、生命科学、物理科学不同参照系，不能停留在广义相对论参照系理论。

太空发展的物理模拟、生命模拟、社会模拟和智力模拟不一定赚钱，但政府应该资助，前沿研究性大学应该投入。其中的物理模拟应该根据生命、社会、智力的需要作针对性研究。

巨量计算力过剩后，正是展开那些模拟好时机，对云平台也是雪里送炭。但需要超前研究和规划。鉴于欧盟旗舰研究Human Brain project连模拟鼠脑、猫脑都失败，各种太空模拟需要新的科学理论基础和人机藕合结构及接口。美国BRAIN initiative项目思路也不行。

万事开头难，要整理出思路不容易。很多人因此放弃思考，求助于人工智能。但人工智能可靠吗，真比吸毒还有效？医生因此诊断力退化，更可怕！

企业想借人工智能之名提升股价。但只有Nvidia能一时暴发。云平台得小利，已被Nvidia超越。OpenAI技术令人失望，无法盈利。应用端一旦断裂，人工智能市场便成庞氏链，硅谷和斯坦福模式也不得不转型。

Nvidia、OpenAI爆发前很多年，我就在批评人工智能问题、倡导科学逻辑和智力新科学研究。是他们想碾压我的科学理论。黄仁勋先生还指望五年内实现通用人工智能AGI、万亿人形机器人市场和人工智能医生？我不研究光刻机和芯片制造，也不参与Nvidia布局Intel之类的活动。

人工智能无法理解自然语言语义和科学理论，也就无法理解生理和疾病。人类医生依赖不可靠的人工智能，自身诊断能力却退化，测试人工智能医生就更困难，所以更可怕。这是熵趋势，比庸医还可怕。很多庸医知道分寸，不会加速病情，或知道自己不行而推荐好医生，或让专家会诊、研究，也可像孙维案在互联网向行家咨询。有些病不治或开些安慰剂就能好。当然好医院知道挑好医生。这些都是高阶逻辑，需要人类智力。所以人类医生判断力退化，后果极为严重。而人工智能医生不懂高阶逻辑，不知分寸。很多其它行业也有类似情况。

面对庞杂训练大数据，计算机无法有效区分真假、对错，也无法化解别人留下的语言陷阱和情绪陷阱，自己还会漏洞百出地胡编乱造。这些跟word-emb edded vector space等等技术和其它构造性数学无法正确表达自然语言语义有关。即使有些具体错误被纠正，结构上root cause无法纠正，更多、更严重问题会产生。更好处理器和训练方法也无法解决。

几十年前就能在计算机里模拟核试验，也有实用Electronic Design Automation (EDA)。数字孪生和工业自动化一直在成长，但增速和取代人力幅度低于预期。人工智能出路在于更有效的人机耦合，研究人类智力的结构和机制需要新的科学逻辑和智力科学。

自动编程在有些逻辑简单的问题上能帮助人类工程师形成思路。这些逻辑简单问题容易测试，自动编程不可靠也没关系。复杂问题如EDA设计在逻辑上却可能很简单而容易测试。逻辑复杂问题测试困难，纠错更困难甚至不可能，自动编程帮不上忙还会添乱甚至留下隐患。人工智能的改进很快趋缓，自动编程也会如此。高昂投入成本下庞氏链无法持久，软件工程师的薪水会向其它行业工程师趋近，看不见的手会起作用。所以说硅谷和斯坦福模式也不得不转型。

说计算机神经网络是物理学在误人子弟。大学研究和教育怎么办？下列重要问题需要研究：

1. 通用人工智能AGI是否可能？研究切入点：荷尔蒙、思维模式切换和演变、如何克服熵趋势。荷尔蒙会引发fight-or-flight等浅层快捷反应模式、影响理性和价值观；而克服熵趋势是思维模式演变和提升智力的关键。

2. 人形机器人和其它人工智能技术的智力上限在哪里，需要什么样的人机耦合结构和接口？

3. 计算机语言模型会制造什么样的诡辩和误导、导致什么样的智力退化？

我把Google Search上AI overview一些测试结果放在《人形机器人的产业陷阱》附录一。这些最简单的测试就能说明计算机语言模型不懂自然语言语义，当然也不懂科学理论。这些跟“知之为知之，不知为不知”等等高阶抽象逻辑问题有关。我把计算机围棋、语言模型、人形机器人当作人工智能

技术的初级、中级、高级实验工具，但不看好其商业应用。

围棋有明确边界，而一般人类智力并非如此。Critique 则需要超越、探索边界并形成更好抽象。所以计算机围棋只能做低级实验工具。而人形机器人可以进一步验证生命智力和机器智力的区别、思维模式的切换和演变以及智力的进化和退化。

我使用 critique 的古典含义（不同于现代中文“批判”），而不是 critical thinking，也不是实用主义教育家杜威使用的 reflective thinking。因为 critique 一词更严谨，可以跟边界、临界点、奇点等等概念构成有效的逻辑体系。Wikipedia 上有“The goal of critical thinking is to form a judgment through the application of rational, skeptical, and unbiased analyses and evaluation.”这在逻辑上会有问题。

我早就提出统计方法对人工智能测试无效。即使有些具体错误被纠正，但结构上的 root cause 无法纠正，更多、更严重的问题会产生。再好的处理器和训练方法也无法解决。那些测试结果说明：现在学术界、产业界采用的人工智能测试方法完全错误。所以需要进一步研究逻辑主义、直觉主义、形式主义三大数学学派的本质区别和各自缺陷以及如何修正和改进。

未对现有脑/神经和智力研究作系统、审慎评估，轻率使用炒作语言把计算机神经网络当物理学是明显的智力退化。用逻辑研究逻辑、用测量研究测量会造成比哥德尔语句更重要、更广适的问题，斯坦福大学“以人为本”人工智能研究院和吴蔡神经研究所知道吗，怎么教学生?

计算机自然语言处理的进展，比我想象的更差。以下是几个层次的情况：

1. 把语音转化成文字。这是较低的听觉层次。我不研究，所以曾以为已基本解决，实际上还差很远。

2. 翻译。我曾以为大部分已解决，实际上也没有。举个简单例子：俄语 Слышу малиновый звон，слышу 是 hear，малиновый 是 raspberry 或颜色 crimson，звон 是铃或铃声。Google Translate 译成 I hear the crimson ringing，当然错误。我俄语非常差，也知道应该是 raspberry。虽然 Google 可以很快纠正已公布的具体问题，但结构上的 root cause 无法纠正。涉及重要的抽象和高阶逻辑，错误会很多。还常有些低级错误，如某些部分内容在翻译中消失了，人类很容易看出来。所以人类智力和计算机智力明显不同。

3. 根据现有信息回答问题。人工智能往往比人类死记硬背还差，不但会胡编乱造，简单换个方式提问往往就会答错、前言不搭后语。

4. 解决新问题。对少数问题有些价值如：计算机围棋有明确边界，蛋白质结构也是高度受限的局部问题。但人类未来可能找到比人工智能更好的理论解法。而绝大部分问题缺乏很好的边界控制，人工智能就不行了。事实上，语言模型技术已露败相。